

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 20, 1999, 15:08:03 : Search time 14.81 seconds

(without alignments)  
535.776 Million cell updates/sec

Title: US-09-126-945-2

Perfect score: 1785

Sequence: 1 MGSASPOLSSVSPSHLLLP.....GIIRKPDISQRLVYQFVHPI 335

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database: A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	294.5	16.5	451	1	Y01520
2	290.5	16.3	478	1	Y01521
3	277	15.5	454	1	R33365
4	270	15.1	452	1	R44556
5	262.5	14.7	581	1	W47238
6	242	13.6	521	1	W47237
7	240.5	13.5	371	1	W57834
8	238.5	13.4	371	1	W50677
9	236	13.2	520	1	R78185
10	236	13.2	470	1	W78399
11	234	13.1	370	1	W60680
12	224	12.5	555	1	W49010
13	222	12.4	462	1	W60167
14	215.5	12.1	155	1	R45451
15	194.5	10.9	548	1	W07700
16	194.5	10.9	543	1	W07702
17	151.5	8.5	272	1	R13444
18	99	5.5	1618	1	R27205
19	99	5.5	1618	1	R60127
20	97	5.4	1757	1	W84351
21	96.5	5.4	1142	1	W81546
22	94.5	5.3	1244	1	W59358
23	91.5	5.1	789	1	W14055
24	90	5.0	2035	1	R57141
25	88.5	5.0	1974	1	W98391
26	87.5	4.9	1711	1	W70506
27	87.5	4.9	1711	1	W70507
28	86.5	4.8	1805	1	R27204
29	86.5	4.8	1805	1	R27204
30	86.5	4.8	1805	1	R27204
31	85.5	4.8	903	1	W77289
32	84.5	4.7	821	1	R35451
33	84.5	4.7	821	1	R35451
34	84	4.7	569	1	R80632
35	83.5	4.7	500	1	W56791
36	83.5	4.7	733	1	W69430
37	83	4.6	677	1	W06083
38	83	4.6	677	1	W25017
39	83	4.6	702	1	W21184
40	82.5	4.6	2629	1	W61348
41	82	4.6	1247	1	R77348
42	82	4.6	405	1	W44933
43	82	4.6	275	1	W83386

44 81.5 4.6 1723 1 W00645  
45 81.5 4.6 1461 1 W64468

Mouse DEC-205. L1g  
Human secreted pro

## ALIGNMENTS

RESULT	ID	Query Match	Score	DB	Length
1	Y01520	standard; Protein; 451 AA.	294.5	1	451
2	Y01521	standard; Protein; 478 AA.	290.5	1	478
3	Y01521	standard; Protein; 478 AA.	290.5	1	478
4	Y01521	standard; Protein; 478 AA.	290.5	1	478
5	Y01521	standard; Protein; 478 AA.	290.5	1	478
6	Y01521	standard; Protein; 478 AA.	290.5	1	478
7	Y01521	standard; Protein; 478 AA.	290.5	1	478
8	Y01521	standard; Protein; 478 AA.	290.5	1	478
9	Y01521	standard; Protein; 478 AA.	290.5	1	478
10	Y01521	standard; Protein; 478 AA.	290.5	1	478
11	Y01521	standard; Protein; 478 AA.	290.5	1	478
12	Y01521	standard; Protein; 478 AA.	290.5	1	478
13	Y01521	standard; Protein; 478 AA.	290.5	1	478
14	Y01521	standard; Protein; 478 AA.	290.5	1	478
15	Y01521	standard; Protein; 478 AA.	290.5	1	478
16	Y01521	standard; Protein; 478 AA.	290.5	1	478
17	Y01521	standard; Protein; 478 AA.	290.5	1	478
18	Y01521	standard; Protein; 478 AA.	290.5	1	478
19	Y01521	standard; Protein; 478 AA.	290.5	1	478
20	Y01521	standard; Protein; 478 AA.	290.5	1	478
21	Y01521	standard; Protein; 478 AA.	290.5	1	478
22	Y01521	standard; Protein; 478 AA.	290.5	1	478
23	Y01521	standard; Protein; 478 AA.	290.5	1	478
24	Y01521	standard; Protein; 478 AA.	290.5	1	478
25	Y01521	standard; Protein; 478 AA.	290.5	1	478
26	Y01521	standard; Protein; 478 AA.	290.5	1	478
27	Y01521	standard; Protein; 478 AA.	290.5	1	478
28	Y01521	standard; Protein; 478 AA.	290.5	1	478
29	Y01521	standard; Protein; 478 AA.	290.5	1	478
30	Y01521	standard; Protein; 478 AA.	290.5	1	478
31	Y01521	standard; Protein; 478 AA.	290.5	1	478
32	Y01521	standard; Protein; 478 AA.	290.5	1	478
33	Y01521	standard; Protein; 478 AA.	290.5	1	478
34	Y01521	standard; Protein; 478 AA.	290.5	1	478
35	Y01521	standard; Protein; 478 AA.	290.5	1	478
36	Y01521	standard; Protein; 478 AA.	290.5	1	478
37	Y01521	standard; Protein; 478 AA.	290.5	1	478
38	Y01521	standard; Protein; 478 AA.	290.5	1	478
39	Y01521	standard; Protein; 478 AA.	290.5	1	478
40	Y01521	standard; Protein; 478 AA.	290.5	1	478
41	Y01521	standard; Protein; 478 AA.	290.5	1	478
42	Y01521	standard; Protein; 478 AA.	290.5	1	478
43	Y01521	standard; Protein; 478 AA.	290.5	1	478

KM Chicken: C-11 protein: cell calcification inhibiting activity;  
 KM cell calcification inhibiting agent; c-erg protein; arthritis deformans;  
 KM ossification; spinal column ligament.  
 OS Gallus sp.  
 PN J11075871-A.  
 PD 23-MAR-1999.  
 PF 29-MAY-1998: 166076.  
 PR 20-JUN-1997: US-050297.  
 PR 18-JUN-1997: US-878177.  
 PA (CHUS) CHUGAI PHARM CO. LTD.  
 PA (UPE-) UNITIV PENNSYLVANIA.  
 DR WPI: 99-257708/22.  
 DR N-PSDB: X26552.  
 PT An active protein for inhibiting cell calcification - useful for  
 PT measuring the calcification of a cell, for diagnosing arthritis  
 PT deformans or ossification of spinal column ligament  
 PS Claim 5: Page 9-10; 15pp: Japanese.  
 CC The present sequence represents chicken c-erg protein. The specification  
 CC also describes a chicken C-11 protein (Y01520) which has cell  
 CC calcification inhibiting activity and a cell calcification inhibiting  
 CC agent containing c-erg protein. The proteins are used for measuring the  
 CC calcification of a cell, for diagnosing arthritis deformans or  
 CC ossification of spinal column ligament.  
 SQ Sequence 478 AA:

Query Match 16.3%; Score 290.5; DB 1: Length 478;  
 Best Local Similarity 26.4%; Pred. No. 4.9e-18;  
 Matches 96: Conservative 41; Mismatches 114; Indels 113; Gaps 13;

```

OY 67 YEDSSMAAKADGASSREPP-----EEPEQCPVIDSQAPAGSLDLV 108
DB 41 YGQTSKMSPRVPOQDMLSQPARVTIKMECNPNQVNGSRNSPDOSVAKGKMSSDNV 100
OY 109 P---GGLTLEESLSFOVQSMVGEVLKQIEPKLNTADPMSPSNVQKMLMTHQ 165
DB 101 GMYNSYSEKTI--PPPMATTN-----RRVYAPDPTLMTSDHVMLEMAVKE 149
OY 166 YRLPMG- KAFQELAGKELCAMESEOFROSP- LGDVLHNL-----DIW 209
DB 150 YGLPDVDILLFQNIQDKELCKMTKDDQRLFPSTYNADILLSHLHRTPLPHLTSDDVD 209
OY 210 K-----SAAMMERSTSPGAIHYCASTS-----EESN----- 236
DB 210 KALONSPLMHAARNTGATFIPTNTSVPEARTQRTTTPRDPDPEQARRSAWTSHTQS 269
OY 236 -----TDEVDSSCSGQPIHLMOFLKELLKPHSYGR 267
DB 270 KATOPSSSTVPKTEDQRPQDLPYQILGPTSSRLANPGSQ- IOLMOFLLE-LLSDSSNSN 327
OY 268 FIRWLNKKGKIFKIDSQVAVRLMGIRKNRPAMNTDKLSRSIRQYKKGIIRKPDISQRL 327
DB 328 CITW-EGTNGEFKMTDPDEVARRMGERKSPNNMYDKLSRLALRYYYDKNIMTKVHPRESS 386
OY 328 VYOF 331
DB 387 MYKY 390

```

RESULT 3  
 R33365 ID R33365 standard; Protein: 454 AA.  
 AC R33365;  
 DT 30-JUN-1993 (first entry)  
 DE Sequence of purine-rich repeat (GA repeat) binding protein  
 DE (GABP) subunit alpha.  
 KM GA binding protein; cis-regulatory element;  
 KM VP16 mediated induction.  
 OS Mus musculus.  
 PN M09304166-A.  
 PD 04-MAR-1993.  
 PF 17-AUG-1992: 006748.  
 PR 16-AUG-1991: US-746032.

PA (CARN-) CARNEGIE INST WASHINGTON.  
 PI Lamarco RL, Mc Knight SL, Thompson CC;  
 DR WPI: 93-093998/11.  
 DR N-PSDB: Q37480.  
 PT DNA encoding GA binding protein sub-unit - allows investigation  
 PT of sub-unit sequence motif functions, for control of rapid cell  
 PT division e.g. in cancer.  
 PS Disclosure: Fig 2a; 68pp: English.  
 CC A cis-regulatory element required for virion associated protein VP16  
 CC mediated induction of herpes simplex virus 1 (HSV1) immediate early  
 CC (IE) genes consists of three imperfect repeats of the purine-rich  
 CC hexanucleotide 5'-CGGAR-3'. A protein complex capable of avid  
 CC interaction with the purine-rich repeats (GA repeats) has been  
 CC identified in soluble preparations of rat liver nuclei. This GA  
 CC binding protein (GABP) consists of two separable subunits.  
 CC Applicants have isolated cDNA clones encoding both subunits of GABP  
 CC and have revealed that one (GABP alpha) is related to the Ets  
 CC transforming protein, while the other (GABP beta) contains a  
 CC series of 33-amino acid repeats related in sequence to a variety  
 CC of proteins.  
 SQ Sequence 454 AA:

Query Match 15.5%; Score 277; DB 1: Length 454;  
 Best Local Similarity 28.8%; Pred. No. 7.4e-17;  
 Matches 83: Conservative 43; Mismatches 116; Indels 46; Gaps 9;

```

OY 78 PGASSREPEPEPEQCPVIDSQAPAGSLDVPGLTLEESLSFOVQSMVGEVLKQIETA 137
DB 126 PDAHHAFAEAHLVEAQL-----TLDGKTHITTSIDSTSQVYRMAA--ALBEYRKE 176
OY 138 CKLLNTADPMSPSNVQKMLMTHEQYRLPPMGKAFQELAGKELCAMESEOFROSP 197
DB 177 QERGLTIDPDIRKSTQVLMVVMVVKKEFSMTDIDLTLNISGRELCINSIOEDFEQRPV 236
OY 198 GGDVLHNLDIRKSAAMMER-----TSPGAIHYCASTS----- 232
DB 236 RGEILMSHELKRYLALSOEQMNEIYITDQVQIIPASVPATYTTIKYINSSAKAK 295
OY 232 -EESWIDSEVDSSCSQ-----PIHMOFLKELLKPHSYGRFIWLNKREKGFIEDS 284
DB 296 VQSPRISGEDRSSPGRNGNNGOIOLOMOFLLELLDIXDARD-CISWYDE-GEFRLNQP 353
OY 285 AQVAVRLMGIRKNRPAMNTDKLSRSIRQYKKGIIRKPDISQRLVYOF 332
DB 354 ELVAQKMGQRKNKPTMYNTERKSLRALRYYYDKNIMTKVHPRESS 400

```

RESULT 4  
 R44556 ID R44556 standard; Protein: 452 AA.  
 AC R44556;  
 DT 26-MAY-1994 (first entry)  
 DE Human HUM-FLI-1 gene product.  
 KM Chromosomal translocation; chimeric; chimaeric; Ewing sarcoma;  
 KM Ews gene; malignant melanoma; hum-fl-1; Clone BM025;  
 KM primitive peripherai neuroectodermal tumour; human chromosome 11;  
 OS Homo sapiens.  
 PN M09323549-A.  
 PD 25-NOV-1993.  
 PF 19-MAY-1993: F00494.  
 PR 20-MAY-1992: FR-006123.  
 PA (CNRS) CNRS CENT NAT RECH SCI.  
 PI Aurias A, Delattre O, Desmaza C, Melot T, Peter M;  
 PI Plougastel B, Thomas G, Zucman J;  
 DR WPI: 93-386580/48.  
 DR N-PSDB: Q50644.  
 PT New nucleic acid of EWS gene and its hybrid(s) - contg. gene  
 PT sequence involved in chromosomal trans-location, also derived  
 PT mRNA, probes, fusion proteins etc., for diagnosis and treatment  
 PT of Ewing sarcoma and melanoma  
 PS Disclosure: Fig 7; 123pp; French.

CC The probe 11R1 was used to screen a human marrow cDNA library  
 CC (Clontech cat. # HL1058). The clone BM025 was identified and  
 CC sequenced. It represents the entire coding region together with  
 CC 5'- and 3'-UTRs of the Hum-fil-1 gene.  
 SQ Sequence 452 AA:

Query Match 15.1%; Score 270; DB 1; Length 452;  
 Best Local Similarity 25.9%; Pred. No. 3,1e-16;  
 Matches 103; Conservative 50; Mismatches 137; Indels 108; Gaps 17;

OY 2 GSASGLSSVSRSLLLPDVSRTGLEKAAGVGLERRDWSPPAPTEGCG---LSA 57  
 DB 3 GIKKALSVSDQSL-----PDSAYGAALHPKADMTAS--GSPDYGPHKINP 50  
 OY 58 FYLSFTFDMLYPDDSSWAKAPGASSRE-----PPEPEQCPVDSQAPAGSLDLYPGG 111  
 DB 51 L-----PPOQEMINQPVRYVNRKREYDHNGSRESVDCSV-----SKCSKLVGGG 95  
 OY 112 LT-----LEHSLDEQVQSMVNGEVLKIDTACKLLMTADPMDWSPSNQKWLWTE 163  
 DB 96 ESNPNMNTSYMDEKNGPPPPNTTNE-----RRYVPPADPLMTQEHYKQLEMAI 146  
 OY 164 HOYRLPPMCKA-FOELACKELCAMSEEOFRORSPL-GGDVLAHL----- 207  
 DB 147 KEYSLEIDTSEFQNMNDGKELCKMKNKEDFLRATTLTYNTEVLLSHLSYLRESSLAYNTTS 206  
 OY 207 -----DIKSKAMMKERISP-----GAIYCASTSEE----- 234  
 DB 207 HTDQSSRLSVKEDPSYDVSRRAGMGNMNSGLNKSPPLEGAGTIRKNEQRPQPPQIL 266  
 OY 234 SWTDEVSQSCSGPILHMQFKELLKPHSYGRFIRMLNKEKGFIEDSAQVRLMG 293  
 DB 267 GPTSSRLANPGSGO-IQLMQLLE-LISDSANASCITW-EGTNGEFKMTDPDEVARRNGE 323  
 OY 294 RKNRPAMNYDKLSRSIROYKKGIRKPDISQRLVYQF 331  
 DB 324 RKSQPNMNYDKLSRLALRYYYDKNIMTKVH-GKRYAYKF 360

## RESULT 5

W47238  
 ID W47238 standard; Protein: 581 AA.  
 AC W47238;  
 DT 22-MAY-1998 (first entry)  
 DE Human lymphoid-specific transcription factor NERF-2.  
 KW ets-related; human; lymphoid-specific transcription factor; NERF-2;  
 KM screening assay.  
 OS Homo sapiens.  
 PN US5721113-A.  
 PD 24-FEB-1998.  
 PF 03-JAN-1995; 368281.  
 PR 03-JAN-1995; US-368281.  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PI Kunsch CA, Libermann TR, Oetting JP;  
 DR WPI: 98-168405/15.  
 DR N-PSDB: V15530.  
 PT NERF-1 and NERF-2 nucleic acids - encode ets-related human  
 PT lymphoid-specific transcription factors  
 PS Disclosure: Columns 27-30; 17pp; English.  
 CC The present sequence is the ets-related lymphoid-specific  
 CC transcription factor NERF-2, which can be used in screening assays  
 CC for drugs that modulate NERF activity, and to treat patients having  
 CC need of NERF-2, ets-related factors are involved in, e.g. cancer  
 CC development, retrovirus replication, T-cell cytokine production,  
 CC cell cycle effects, growth regulation and cell differentiation.  
 SQ Sequence 581 AA;

Query Match 14.7%; Score 262.5; DB 1; Length 581;  
 Best Local Similarity 25.2%; Pred. No. 2,1e-15;  
 Matches 83; Conservative 52; Mismatches 108; Indels 87; Gaps 10;

OY 26 TGLEKAAGVGLERRD-----WSPSPATPEOGLSAFYLFDMYL-PEDSSWAA 75  
 DB 11 TLEISSNGVNOESDSVSTYPAVIEVPASALBEGYAAQVGYDDETYLMQDVA--- 68  
 OY 76 KAPGASSREPEPEPEQPVIDS-----QAPAGSLDLYPGGLTLEHSLQ 121  
 DB 68 -----EEOEVEENENETVEASVHSSNAHCTDRTIEAABALLHMSPTCLDRSRPEF 119  
 OY 122 VQSMVNGEVLKIDTACKLLMTADPMDWSPSNQKWLWTEHQRRLPMPKAOELAGK 181  
 DB 120 IHAAAMPDVI---TETVEVSTEESEPMDTSP-----IPSPDSHEPKKK 162  
 OY 182 ELCAWSEOFORSPLGGDVLHAHLDIWKSAAWKKERTSPGAIHYCASTSESWTDEVD 241  
 DB 163 K---VGRKPKTQOSPISNG-----SPELGIKKKPRG----- 192  
 OY 242 SSCSGOPILHMQFKELLKPHSYGRFIRMLNKEKGFIEDSAQVRLMGIRNRPMN 301  
 DB 192 ---KGNTTYLWEFLDLQDKNTCPRYIKWQREKGFILVDSKAASKLMGKHKKPDMN 248  
 OY 302 YDKLSRSIROYKKGIRKPDISQRLVYQF 331  
 DB 249 YETMGRLRITYQRIKLAKE-GQRLVYQF 277

## RESULT 6

W47237  
 ID W47237 standard; Protein: 521 AA.  
 AC W47237;  
 DT 22-MAY-1998 (first entry)  
 DE Human lymphoid-specific transcription factor NERF-1.  
 KW ets-related; human; lymphoid-specific transcription factor; NERF-1;  
 KM screening assay.  
 OS Homo sapiens.  
 PN US5721113-A.  
 PD 24-FEB-1998.  
 PF 03-JAN-1995; 368281.  
 PR 03-JAN-1995; US-368281.  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PI Kunsch CA, Libermann TR, Oetting JP;  
 DR WPI: 98-168405/15.  
 DR N-PSDB: V15529.  
 PT NERF-1 and NERF-2 nucleic acids - encode ets-related human  
 PT lymphoid-specific transcription factors  
 PS Disclosure: Columns 21-24; 17pp; English.  
 CC The present sequence is the ets-related lymphoid-specific  
 CC transcription factor NERF-1, which can be used in screening assays  
 CC for drugs that modulate NERF activity, and to treat patients having  
 CC need of NERF-1, ets-related factors are involved in, e.g. cancer  
 CC development, retrovirus replication, T-cell cytokine production,  
 CC cell cycle effects, growth regulation and cell differentiation.  
 SQ Sequence 521 AA;

Query Match 13.6%; Score 242; DB 1; Length 521;  
 Best Local Similarity 25.3%; Pred. No. 1,2e-13;  
 Matches 66; Conservative 45; Mismatches 72; Indels 78; Gaps 8;

OY 99 QAPAGSLDLYPGGLTLEHS-----LEQVQSMVNGE---VLKD----- 134  
 DB 7 EGPFTQDILLRAVEASVHSSNAHCTDRTIEAABALLHMSPTCLDRSRPEFIHAAMP 66  
 OY 134 ---LETACKLLNTRADPMDWSPSNQKWLWTEHQRRLPMPKAFQELACKAMSEEO 190  
 DB 67 DVITETVEVSTEESEPMDTSP-----IPSPDSHEPKKK---VGRK 108  
 OY 191 FROSPPLGCDVLHAHLDIWKSAAWKKERTSPGAIHYCASTSESWTDEVSQGPPIH 250  
 DB 109 KTOOSPISNG-----SPELGIKKKPRG-----KGNTTY 137  
 OY 251 LMQFLKELLKPHSYGRFIRMLNKEKGFIEDSAQVRLMGIRKRPAMNYDKLSRSIR 310  
 DB 138 LMWFLDLQDKNTCPRYIKWQREKGFILVDSKAASKLMGKHKKPDMNYETMGRLR 197



DB 45 SNPKSLEKTEKASWGE-----QPFMSKTQVLDWISQYENKNKYDAS 88  
 OY 170 PMKAFQELAGKELCAMSEEDFR-QRSLPGDVLAHL-DIMKSA-----AMW----- 216  
 DB 89 AIDFERCDDMGATLNCALDELRLVFGPL-GQQLHAQLRLDSSSSDELSTWILLEKDG 147  
 OY 216 ---KERTSPGAIHYCASTSEESTDEYDS-----SCS----- 246  
 DB 148 MAFOALDPPGPRDQSPFAQELLDGQOASPYHPSGCCAGAPSPSSDVYTAGTASRSS 207  
 OY 246 -----GQP----- 249  
 DB 208 HSSDGSQVDLPDGLKLPESDGFROCKKDPKHKRRGRPRLSKEYMDCLEGRKSK 267  
 OY 249 -----IHLMOFLKELLRLPHSYGRIRLWLNKKGIFKIEDSAQVARKNGIRKNRPAMNYD 303  
 DB 268 HAPRGCTHMEFRLDILHPELNEGMLKWNENHEGVKFLRSEAVAQMLGQKKNSMTYE 327  
 OY 304 KLSRSIRQYKKGIIRKPDISORLYQVF 331  
 DB 328 KLSRAMRYTKREILERVD-GRRLYKVF 354

RESULT 9  
 R78185  
 ID R78185 standard; Protein: 520 AA.  
 AC R78185;  
 DT 09-FEB-1996 (first entry)  
 DE Protein sequence of PEA3-Beta -an ETS transcription factor.  
 KW Transcription factor; Probe: reverse transcription; PCR: primer;  
 KW expression vector; E.coli; COS cell; ras; cancer cell multiplication;  
 KW polyoma virus; transformation.  
 OS Homo sapiens.  
 FH Key  
 FT misc\_difference 84 /note= "encoded by CTC"  
 FT misc\_difference 126 /note= "encoded by AAC"  
 FT misc\_difference 500 /note= "encoded by ACC"  
 FT J07145197-A.  
 PN 06-JUN-1993.  
 PD 25-NOV-1993; 295393.  
 PR 25-NOV-1993; JP-295393.  
 PA (EISA) HIRANO T.  
 PA (HIRA) HIRANO T.  
 DR MPI: 95-237197/31.  
 DR N-PSDB: 091769.  
 PT ETS transcription factor activated by ras - may be used in the study  
 PT of cancer cell proliferation and the proliferation of the polyoma  
 PT virus  
 PS Claim 1: Page 7-9; 9pp; Japanese.  
 CC The amino acid sequence of the novel ETS transcription factor family  
 CC member - PEA3-beta. The gene was isolated from a HepG2 cell line cDNA  
 CC library. The probe for the screening was prepared by reverse  
 CC transcription on HepG2 mRNA followed by PCR using primers 091770-1,  
 CC to produce a probe of 170-200 bp. The screening isolated the full  
 CC length sequence of the transcription factor. The gene was inserted into  
 CC the expression vectors pBluescript KS and pCDV1 for expression of the  
 CC protein in E.coli and COS7 cells, respectively. The ETS transcription  
 CC factor has specificity for and is activated by ras. It is useful as a  
 CC reagent in studies for the elucidation of the mechanism of cancer cell  
 CC multiplication or polyoma virus transformation of cells.  
 SQ Sequence 520 AA;

Query Match 13.2%; Score 236; DB 1; Length 520;  
 Best Local Similarity 25.7%; Pred. No. 4,3e-13;  
 Matches 98; Conservative 48; Mismatches 115; Indels 120; Gaps 18;  
 OY 5 SFGGLSSVSPSH---LLLPDVTYSRTGLEKAAAGVLEGRDMSPPRPATPEGLSAPFLS 61  
 DB 135 TPTPTPLSPSTHONPLPFPPOATLPTSGHAPAGPV--QGVGPAPAPHSLEPGE----- 186

OY 62 YFDMLYEDSSMAKAPGASREPEPECCPVY--DSQAPAGSLDVPGLTLTEHS 119  
 DB 186 -----FOOQTFAY-----PRPHQPLQMPKMMENQYPS-----EDRFQ 219  
 OY 120 EOV-----QSMVGVLEKLDIETACKLNLITADPMDSPSNVQKWLMTHEOYRLP 170  
 DB 220 RQLSEPCHPFPQPGVGDNDNPSYHRQMSERIVPAAP--PPQGFQKEYHDLPLXHEGVPG 277  
 OY 170 ---PMKAFQELAG-----KELCAMSEEDFRQSRPLGQDVLAHLDIMKSAAMKERTSP 221  
 DB 278 MCGPPAHGFQSPMGIRKQEPBDYCVDSVPPNQSSTMGKGYF----- 319  
 OY 222 GAHYCASTSEESTWDS-----VDSC-----SGOP-----IHLW 252  
 DB 319 -----SSHEGSGSYEKDPLLYFDITCVYPERLEGVYKQEPYTRKGPPIQGRSGSLW 371  
 OY 253 QFLKELLKPHSYGRFIRLWLNKKGIFKIEDSAQVARKNGIRKNRPAMNYDKLSRSIRQ 311  
 DB 372 QFLVTLDDP-ANAHFIAMTG--RGMEKLEPEVARRMGIQKNRPAMNYDKLSRSIRY 428  
 OY 312 YKKGIIIRKPDISORLYQVF 332  
 DB 429 YEKGIQKV-AGERYVYKVF 448

RESULT 10  
 W78399  
 ID W78399 standard; Protein: 470 AA.  
 AC W78399;  
 DT 11-MAY-1999 (first entry)  
 DE Rat Ets-2 protein.  
 KW Epsilon-subunit; muscle nicotinic acetylcholine receptor; nAChR; rat;  
 KW neuromuscular junction; transynaptic expression disorder; spinal cord;  
 KW myasthenia gravis; familial infantile myasthenia; acetylcholinesterase;  
 KW slow channel syndrome; degenerative disease; peripheral nerve; stroke;  
 KW skeletal muscle tissue; brain tissue; Alzheimer's disease;  
 KW brain aneurysm; infection; tumour; bleeding; blood clot.  
 OS Rattus norvegicus.  
 PN W09855611-A1.  
 PD 10-DEC-1998.  
 PR 05-JUN-1998; U11695.  
 PR 06-JUN-1997; US-048847.  
 PA (UNMT) UNIV MICHIGAN.  
 PA Goldman D, Sapru MK.  
 DR MPI: 99-059911/05.  
 DR WPI: 99-059911/05.  
 PT New isolated neuromuscular response element - used to develop products  
 PT for treating e.g. neurological or muscular degenerative disease or  
 PT damage to spinal cord, peripheral nerve, skeletal muscle tissue or  
 PT brain tissue  
 PS Claim 32; Fig 4; 69pp; English.  
 CC This sequence represents the rat Ets-2 protein which is an example of  
 CC a neuromuscular response element (NRE) binding protein. NRE corresponds  
 CC to nucleotides -55 to -69 (starting from the transcription start site)  
 CC of the gene for the epsilon-subunit of muscle nicotinic acetylcholine  
 CC receptor (nAChR) in a rat (see x16979). The NRE can be used for  
 CC expression of a heterologous gene at a neuromuscular junction in a  
 CC subject having a transynaptic expression disorder, e.g. myasthenia  
 CC gravis, familial infantile myasthenia, acetylcholinesterase deficiency,  
 CC slow channel syndrome or AChR deficiency, neurological degenerative  
 CC disease, a muscular degenerative disease, damage to a spinal cord,  
 CC peripheral nerve, skeletal muscle tissue or brain tissue, e.g. in  
 CC Alzheimer's disease, stroke, brain aneurysm, brain infection, brain  
 CC tumour, brain bleeding or brain blood clot.  
 SQ Sequence 470 AA;

Query Match 13.2%; Score 236; DB 1; Length 470;  
 Best Local Similarity 21.0%; Pred. No. 3,7e-13;  
 Matches 86; Conservative 44; Mismatches 105; Indels 174; Gaps 12;  
 OY 95 VIDSOAPAGSLDVPGLTLTEHS-----LEOVQSMVGVLEK-----DIETACKLNL 142

DB 39 VLPSISEQTLQEVPTGLDVSVDATCELPPLTPCSKAVNSQALKATESGFHKEQRRLG 98  
OY 143 IRADPMDSVSPVOKMLMTETHOYRLPRGKAFQELAGELCAMESEDFRQNSP-LGGDV 201  
DB 99 IFRNMLNNEQVCCMLHATNEFSLVNVLORFGNNGOMLNLKREPLELAPDFVBDI 158  
OY 202 LHAHLN-----IW-----KSAWKE--- 218  
DB 159 LMEHLEOMIKENOKEDEDOYEENSHLNAPVHINSENTLGFEVGAQPYGMQAPSYLAKDLL 218  
OY 218 -----RTSPGAHYCASTSE----- 233  
DB 219 DGMCPSPATPAALGSEQLQMLPKSRLLNTVSNYCSISQDPFGNMLNLSSSGKPKEDH 278  
OY 233 -----ESWTD-----SEVDSGS----- 246  
DB 279 SPENGDSFESSDSLRRWNSOSSLLDVRVPSFESFEEDCSQSLCKSPMSFEDYIOE 338  
OY 246 -----GQPIHLMOPLKELLRLPHSYGRFRIMLNKEGIFKIED 283  
DB 339 RSDPVEQCKPVIPAAVLGFTGSGPIQLMOFLE-LISDKSCQSPISWTG-DGMFELAD 396  
OY 284 SAQVARIWGIRKRRPAMVYDKLSRIROYKKGIIRKFDISQRLVYQV 332  
DB 337 PDEVARRWGKRKKPKMNEYKLSRGLRYRYDKNIHKTG-GKRYRYREV 444

## RESULT 11

W60680 standard; Protein: 370 AA.

AC W60680: (first entry)

DE Murine transcription factor ESX protein.

KW ESX; epithelial-restricted with serine box; transcription factor;

KW ETS; mouse; epithelial cancer; breast cancer; adenocarcinoma;

KW diagnosis; therapy.

OS Mus sp.

FT Key Location/Qualifiers

FT Region 1..62 "first variable region"

FT Domain 63..102

FT Region 103..187 "A-region/poinded domain"

FT Domain 129..159 "second variable region (claim 1)"

FT Region 144 "transactivation domain"

FT Misc-difference 144

FT Region 188..238 "residue critical for transactivation"

FT Region 239..273 "serine-rich box"

FT Domain 274..354 "third variable region"

FT Region 355..371 "ETS DNA binding domain"

FT Region /note- "fourth variable domain"

PN W09823782-A2.

PD 04-JUN-1998.

PF 26-NOV-1997: U21865.

PR 23-NOV-1997: US-031504.

PA (REGC ) UNIV CALIFORNIA.

PI Benz CC, Chang C, Scott GK;

DR WPI: 98-322755/28.

N-PSDB: V38365.

PT Nucleic acid encoding the ESX transcription activator -

PT over-expressed in epithelial, especially breast, cancers, also

PT related polypeptide(s), antibodies, vectors and transformed cells,

PT useful for diagnosis and treatment of cancer

PS Disclosure: Fig 5: 120pp: English.

CC This polypeptide is encoded by the newly isolated mouse ETS gene

CC (see V38365). It is a novel transcription factor associated with

CC the aetiology of cancers, including epithelial cancers, and is

CC designated ETS for epithelial-restricted with serine box. Human  
CC ESX (see W60677) is an early indicator of breast cancer, being  
CC expressed at the ductal carcinoma in situ stage. Dysfunction of  
CC the ESX gene that results in increased expression is indicative of  
CC epithelial cancer (especially breast cancer but also gastric,  
CC ovarian and lung adenocarcinoma) and of unfavourable prognosis.  
CC Anti-ESX antibodies, antisense nucleic acids or inactive ESX mutant  
CC proteins that inhibit activity of ESX can be used to inhibit growth  
CC and proliferation of neoplastic cells, particularly cancer. ESX  
CC polypeptides, polynucleotides and antibodies are also used for  
CC for diagnosis and monitoring of cancer, and to screen for specific  
CC modulators of ESX, especially potential anticancer agents.  
CC Sequence 370 AA:

Query Match 13.1%; Score 234; DB 1; Length 370;  
Best Local Similarity 21.4%; Pred. No. 46-13;  
Matches 83; Conservative 60; Mismatches 93; Indels 152; Gaps 14;

OY 55 LSAPFLSYFPMKLPEDSSNAKAPGASSREPEPEOCPIVDSONAPAGSLDLPVGLTL 114  
DB 7 ISNVFSYFNAMTSSSDPTLAPAPPT-----FGTEDLV---LTL 43  
OY 115 EHSLEQVOSMVVGEVLKDIETACKLLNITADPMQSPSNVOKMLM--TEHOYRLPMW 172  
DB 44 NNOQM-----LLEGEKASWTSERRQFQSKTOVENISYOYEKKNKIDASSID 90  
OY 173 KAPOELAGELCAMESEQR-ORSPUGDVLNHL-DIKSA-----AWKE----- 218  
DB 91 FSRGNMDGATLSCALEBELRYVGPL-GDOLHAQLDLTSSNDELSTWIELLEKQNSF 149  
OY 218 -----RTSPGAHYCASTSE----- 230  
DB 150 QESLGLGSDGSPRAQELLDGQKASP---YCSYIGFAPSGSSDVSATGATPPOS 206  
OY 230 --TSESWTDSEVDSSCS-----GQP----- 249  
DB 207 SHASDGGSDVDLDLTHESKYFPDDETDYKQKPKRKRPRRLSKYWCLEGGKSK 266  
OY 249 -----IHLMOPLKELLRLPHSYGRFRIMLNKEGIFKIDSQVARIWGIRKRRPAMVY 303  
DB 267 HAPRGTHLMEFINDILIHPELNEGILMKMENRHEGVFKLSRAVAQLMGOKKNSNMTYE 326  
OY 304 KLSRSIRQYKKGIIRKFDISQRLVYQF 331  
DB 327 KLSRAMRYKKRELLERVD-GRRLYYKF 353

## RESULT 12

W49010 standard; Protein: 555 AA.

AC W49010: (first entry)

DE 20-OCT-1998

KW Polyomavirus PEA3 protein;

KW Polyomavirus enhancer activator; PEA3; tumour; suppressor; inhibitor;

KW transformation; HER-2; neu promoter; metastasis; cancer.

OS Polyomavirus.

PN W09830585-A2.

PD 16-JUL-1998.

PF 12-JAN-1998: U00880.

PR 10-JAN-1997: US-780835.

PA (TEXA ) UNIV TEXAS SYSTEM.

PI Hung M, Xing X;

DR WPI: 98-399061/34.

N-PSDB: V32688.

PT Method for repressing transformation of cells - by contacting cell

PT with polyomavirus enhancer activator 3, useful for, e.g. creating

PT or preventing cancer, tumorigenesis and metastasis

PS Disclosure: Page 71-72: 83pp: English.

CC This sequence represents a polyomavirus enhancer activator, PEA3. This

CC sequence is used in a method for repressing transformation in a cell

CC which involves contacting the cell with PEA3 to inhibit a transformed

CC phenotype. This sequence can also be used in a method to suppress the

CC growth of a tumour in a mammal comprising introducing to the mammal a  
 CC PE3A-encoding nucleic acid where the expression of PE3A in the mammal  
 CC results in a decrease in the growth rate of the tumour. PE3A can regulate  
 CC the HER-2/neu promoter by transcriptional repression and acts as a tumour  
 CC repressor. The PE3A can be used for reducing a transforming,  
 CC tumorigenic or metastatic potential of a cell. It can be used for the  
 CC prevention and treatment of such transformation-driven events as cancer,  
 CC tumorigenesis and metastasis.  
 SO Sequence 555 AA:

Query Match 12.5%; Score 224; DB 1; Length 555;  
 Best Local Similarity 23.8%; Pred. No. 5.6e-12;  
 Matches 117; Conservative 38; Mismatches 128; Indels 208; Gaps 24;

```

OY 12 SPSHLLPPDVSR-----GLEKAAAGVLERR-----42
DB 40 SPCHAEHPAAAPQOTGPQVSASARGPGVAGSGGHEERMRKGYLDQRYVPTFCSKSPG 99
OY 42 -----DNSSPPATPE-----QGLSAFYLSYFDMLYEDSSMAKA-----78
DB 100 NSSLGELALVPOGKLMPSLPPSDESLFDLSHF-----QETWLAAGVDPDS 148
OY 78 -----PGASSREPP-----DEPQC-----PVID 97
DB 149 DEGFVDFHSENLAHSPTRIKKEPSPRTDPAALSCSRPLRYNHGOCXLRQIAIK 208
OY 98 SQAPAGSLDLVPGGLTLEHS-LEQVQSNVGEVLKDIETACKL---NITADPMQSP 152
DB 209 SPAPG-----APGQSPLOPFSRAEQOSLRLASSSSQSHRGHGLDEHSSVFQOPVDMC- 263
OY 153 SNVQKMLLTHEQRYR-----PMGKAFQELAGKELCAMESEQFR-----195
DB 263 -----HSTSPQGGREPLPAPYOHLSPPCPYPOONTQOEYHDPYLEDAGCP 311
OY 195 -SPLGVDLVAHLWDIKSAMNKERTS-----PGAHYCAST-----SEESWTDE 239
DB 312 ASSQGG--VSGHRYPGAGVYIKQERTDFAVSDVPG---CASMYLHPGFGSPSPGDV 365
OY 240 V-----DSSC-----SGCP-----IHLMOFLKELLRP 262
DB 366 MCGYKSLRPPEDDVCIYKKEGDIKQEGIGAFREGPYORGLQIOMQFLVALDPP 425
OY 263 HSYGRFIRMLNKEKGI-FRIEDSAOYARLMGIRKNRPAMNYDKLSISROYKKGIIRP 321
DB 426 -TNAHFIATG--KGMFELIEPEVARLMGIRKNRPAMNYDKLSISLRYEKGIMQV 482
OY 322 DISQRLVYQFV 332
DB 483 -AGERVYVKFV 492

```

RESULT 13  
 W00167  
 ID W00167 standard; Protein: 462 AA.  
 AC W00167;  
 DT 25-APR-1997 (first entry)  
 DE ELAF matrix metalloproteinase regulator.  
 KW ELAF; matrix; metalloproteinase; regulator; infiltration; cancer;  
 KM metastasis; cell; control; antisense; decoy; DNA binding region;  
 KW target DNA; ribosome; induction; diagnosis; detection; treatment;  
 KM mammary cancer; fibrosarcoma; osteosarcoma; lung cancer.  
 OS Homo sapiens.  
 FH Key  
 FT domain  
 FT 315..399  
 FT /note= "DNA binding domain"  
 FT region 126..222  
 FT /note= "glutamine rich region"  
 PN W09624379-A1.  
 PD 15-AUG-1996.  
 PF 09-JAN-1996; J00016.  
 PR 08-FEB-1995; JP-020173.  
 PA (TAKI ) TAKARA SHUZO CO LTD.

PI Fujinaga K, Higashino F, Yoshida K:  
 DR WPI: 96-384227/38.  
 DR N-PSDB: T37087.  
 PT Control of cancer cell infiltration by ELAF gene expression  
 PT regulation - also diagnosis of cancer by detection of ELAF gene  
 PT expression products  
 PS Example 5: Pages 38-42; 65pp; Japanese.  
 CC The present sequence is the ELAF protein, which is a matrix  
 CC metalloproteinase regulator. The infiltration and metastasis of  
 CC cancer cells can be controlled by regulating the expression and  
 CC expression products of the ELAF gene. This may be accomplished by  
 CC inducing antisense DNA or RNA for the ELAF gene, a decoy gene  
 CC expressing the DNA binding region of the ELAF protein, the target  
 CC DNA for the DNA binding region of the ELAF protein or ribosomes  
 CC corresponding to the ELAF gene mRNA. Cancer can be diagnosed by  
 CC detecting ELAF gene expression products, e.g. ELAF protein or mRNA.  
 CC These methods may be used in the treatment and diagnosis of cancer,  
 CC e.g. mammary cancer, fibrosarcoma, osteosarcoma, lung cancer, etc.  
 SO Sequence 462 AA:

Query Match 12.4%; Score 222; DB 1; Length 462;  
 Best Local Similarity 27.3%; Pred. No. 6.5e-12;  
 Matches 96; Conservative 37; Mismatches 115; Indels 104; Gaps 17;

```

OY 44 SPSPATPEQGLSAFYLSYFDMLYEDSSMAKARGA--SSREPPREEPQCPVIDSQAP 101
DB 89 SKRPPLRYNHGOCXLSSTYDP--PROAIRSPARGALQSPLOPPRAEQONTFRS---144
OY 102 AGSLDLVPGGLTLEHSLQVQSNVGEVLKDIETACKLNTIADPMQSPSNVQKMLW 161
DB 144 SGTSPHPHGHGLGHS-----SVFQOPDICHG-----F 173
OY 162 TEOYRLRPMGKAFQELAGKELCAMESEQFR--RSPV-----GCVLVAHLADI 208
DB 174 TQGGGREDLPAPYOHLSPPCPYPOQSFQOEYHDPYLEDAGAPVADGG--VNGHRYP 231
OY 209 WSAAMNKERT-----SP--GAHY-----CA- 229
DB 222 GAGVYIKQERTDFATDSVTCASMYLHTEGSGSPGDMGNGYERPLRFPDDVCV 291
OY 229 -STSESWTDESVDSGSGP-----IHLMOFLKELLKPSYGRFIRMLNKEKGI-FK 280
DB 292 PKFEGDIEQEGVAFREGPYORGLQIOMQFLVALDPP-TNAHFIATG--RGMEERK 348
OY 281 IEDSAOYARLMGIRKNRPAMNYDKLSISROYKKGIIRKPIISQRLVYQFV 332
DB 349 LIEPEVARLMGIRKNRPAMNYDKLSISLRYEKGIMQV-AGERVYVKFV 399

```

RESULT 14  
 R45451  
 ID R45451 standard; Protein: 155 AA.  
 AC R45451;  
 DT 11-JUL-1994 (first entry)  
 DE Adenovirus ELA-F protein.  
 KW Adenovirus; cancer; ets oncogene; HeLa cell; enhancer core sequence;  
 KM methylation.  
 OS Human adenovirus.  
 PN J05328975-A.  
 PD 14-DEC-1993.  
 PF 02-JUN-1992; 165453.  
 PR 02-JUN-1992; JP-165453.  
 PA (TAKI ) TAKARA SHUZO CO LTD.  
 PA WPI: 94-021923/03.  
 DR N-PSDB: 055149.  
 DR Novel ELA-F gene - for production of adenovirus ELA-F and cancer  
 PT research  
 PS Claim 1: Page 5; 7pp; Japanese.  
 CC A clone comprising the adenovirus ELA-F gene was isolated by  
 CC screening a HeLa cell cDNA library. The cDNA insert from the clone  
 CC contains a 473bp open reading frame which codes for a protein  
 CC having sequence R45451.

SQ Sequence 155 AA:

Job time: 516 sec

## Query Match

Best Local Similarity 12.1%; Score 215.5; DB 1; Length 155;  
 Matches 47; Conservative 14; Mismatches 19; Indels 5; Gaps 4;

QY 249 IHMOFLKELLKPHSYGRFIRMLNKEGI-FRIEDSAQVARIWGIRKRPAMNDKLSR 307

DB 12 LQLOMFLVALLDDP-TNAHFIAWTG--KGMFKLIEPEEARLWGIRKRPAMNDKLSR 68

QY 308 SIROYKKGIIRKPDISQRLVYQFV 332

DB 69 SLRYEYKGIKMKV-AGERVYKVF 92

## RESULT 15

W07700 W07700 standard; Protein; 548 AA.

AC W07700:

DT 06-APR-1997 (first entry)

DE Human ETS2 repressor factor (ERF).

KM ETS2 repressor factor; ERF; transcriptional repressor;

KW tumour suppressor; tumour; cancer; oncoprotein; therapy.

OS Homo sapiens.

FH Key Location/Qualifiers

FT domain 29..106

FT /label= "DNA-binding\_domain"

FT /note= "ets-like DNA binding domain"

FT 472..530

FT /label= "Active\_repressor\_domain"

FT /note= "(Claim 20)"

FT domain

FT W09639517-A1.

PD 12-DEC-1996.

PF 04-JUN-1996: 010177.

PR 05-JUN-1995: US-469412.

PA (USSH ) US DEPT HEALTH &amp; HUMAN SERVICES.

PI Athanasiou MA, Beal GJ, Blair DG, Fisher RJ, Mayrothalassitis GJ;

PI Sgouras D N;

DR WPI: 97-043139/04.

DR N-PSDB: T47198.

PT New DNA encoding ETS2 repressor factor - useful for reducing

PT tumorigenicity, esp. oncogene associated tumour cells

PS Claim 1; Page 59-61; 101pp; English.

CC Novel human ETS2 repressor factor (ERF) (W07700) is the first member

CC of the ETS family to be identified as a transcriptional repressor in

CC mammalian cells. Its amino acid sequence was deduced from a cDNA

CC clone (T47198) derived from K562 cells. ERF and alternatively

CC spliced ERF (see also W07701) show no homology to other known

CC proteins. The ERF repressor domain in combination with a

CC heterologous transcription factor having a binding domain can

CC be used as novel transcriptional repressors to reduce

CC tumorigenicity associated with inappropriate expression of the

CC GAL4, NFkappaB (HIV), MYC (Burkitt lymphoma), Fli-1 (Ewing's

CC sarcoma) and Esr1 transcription factors.

SQ Sequence 548 AA:

Query Match 10.9%; Score 194.5; DB 1; Length 548;

Best Local Similarity 47.3%; Pred. No. 2.5e-09;

Matches 43; Conservative 13; Mismatches 32; Indels 3; Gaps 3;

QY 241 DSSCGGPIHMOFLKELLKPHSYGRFIRMLNKEGI-FRIEDSAQVARIWGIRKRPAM 300

DB 19 ESSPESRQIOLMHLFIELRLK-EEYGVIAW-QGDYGEFVIRKDPDEVARLWGVRKCRPQM 76

QY 301 NYDKLSRSIRROYKKGIIRKPDISQRLVYQF 331

DB 77 NYDKLSRALRYRNKRLAKTK-GRRTYKF 106

Search completed: November 20, 1999, 15:16:39